



PCT09

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:10

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

3 <110> APPLICANT: Nielsen, Jens
 4 Nissen, Torben L
 5 Nielland-Brendt, Morten C
 7 <100> TITLE OF INVENTION: Metabolically engineered microbial cell with an altered
 8 metabolite production
 10 <140> FILE REFERENCE: NIELSEN=4
 12 <140> CURRENT APPLICATION NUMBER: 09/743,414A
 13 <141> CURRENT FILING DATE: 2001-01-10
 15 <150> PRIOR APPLICATION NUMBER: PCT/DK99/00397
 16 <151> PRIOR FILING DATE: 1999-07-12
 18 <160> PRIOR APPLICATION NUMBER: PA 149: 00967
 19 <161> PRIOR FILING DATE: 1999-07-10
 21 <160> NUMBER OF SEQ ID NOS: 14
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <210> LENGTH: 1395
 27 <210> TYPE: DNA
 28 <210> ORGANISM: Azotobacter vinelandii
 30 <210> FEATURE:
 31 <211> NAME/KEY: CDS
 32 <212> LOCATION: (1)..(1395)
 33 <213> OTHER INFORMATION:

W--> 35 <400> 1

36 atg gct gta tat aas tac gat gtg gtg gta atc ggc aca ggc cct gct 48
 W--> 37 Met Ala Val Tyr Xaa Tyr Asp Val Val Val Ile Gly Thr Gly Pro Ala
 38 1 5 10 15
 40 ggc aaa ggg gaa ggg atg aat gcc gtg aag gcc ggg cgc aag gta ggc 96
 41 Gly Glu Gly Ala Ala Met Asn Ala Val Lys Ala Gly Arg Lys Val Ala
 42 20 25 30
 44 gtc ctg gat gat cgc ccc cag gtc ggc ggc aac tgc acc cac ctg gga 144
 45 Val Val Asp Asp Arg Pro Gln Val Gly Gly Asn Cys Thr His Leu Gly
 46 35 40 45
 48 aac att ccc tcc aag ggc ctg cgc cac tgg gtg cgg cag atc atg cag 192
 49 Thr Ile Pro Ser Lys Ala Leu Arg His Ser Val Arg Gln Ile Met Gln
 50 50 55 60
 51 ttc aac aac aat cag ctg ttc cgc cag atc ggc gag cgc cgc tgg ttt 240
 52 Tyr Asn Asn Asn Pro Leu Phe Arg Gln Ile Gly Glu Pro Arg Trp Phe
 53 65 70 75 80
 55 tcc ttc gcc gat gtc ctg aag agc gcc gag cag gtc atc gcc aag cag 288
 56 Ser Phe Ala Asp Val Leu Lys Ser Ala Glu Gln Val Ile Ala Lys Gln
 57 85 90 95
 59 gtc tcc tcc cgc acc ccc tac tat ggc cgc aac tgc atc gat acc ttc 336
 60 Val Ser Ser Arg Thr Gly Tyr Tyr Ala Arg Asn Arg Ile Asp Thr Phe

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,414A

DATE: 01/31/2003

TIME: 17:52:17

Input Set : A:\NIELSEN4.txt

Output Set : N:\CRF4\01312003\I743414A.raw

```

61      100      105      110
64 ttc ggg acc gag agc ttc ttc gac gar cac acc atc gag gtc gtc cgc 384
65 Phe Gly Thr Ala Ser Phe Cys Asp Glu His Thr Ile Glu Val Val His
66      115      120      125
68 ctg aac ggt atg gtc gaa aag ctg gtg gcc aag cag ttc gtc atc gag 432
69 Leu Asn Gly Met Val Glu Thr Leu Val Ala Lys Gln Phe Val Ile Ala
70      130      135      140
71 acc gga tgg cgt cca tac ccc ccc gcc gat gtc gat ttc acc cat cag 480
72 Thr Gly Ser Arg Pro Tyr Arg Pro Ala Asp Val Asp Phe Thr His Pro
73      145      150      155      160
74 cgg atc tac gac agc gag acc acc ctg agc ctg ggc cac acc ccc ctc 528
75 Arg Ile Tyr Asp Ser Asp Thr Ile Leu Ser Leu Gly His Thr Pro Arg
76      165      170      175
78 cgg ttc atc atc tac gga gag ggg gtg atc ggc tgc gaa tat gcc ttc 576
79 Arg Leu Ile Ile Tyr Gly Ala Gly Val Ile Gly Cys Glu Tyr Ala Ser
80      180      185      190
84 atc ttc agt gga ctg ggt gtg ctg gtc gac ctg atc gac acc cgc gac 624
85 Ile Phe Ser Gly Leu Gly Val Leu Val Asp Leu Ile Asp Asn Arg Arg
86      195      200      205
88 cac ctg ctg aat ttc ctg gac gag gaa atc tcc gag tgg ctg agc ttc 672
89 Gln Leu Leu Ser Phe Leu Asp Asp Glu Ile Ser Asp Ser Leu Ser Tyr
90      210      215      220
92 cac ctg cgc aac aac aac gtg ctg atc cgc cac aac gar gaa tac gag 720
93 His Leu Arg Asn Asn Asn Val Leu Ile Arg His Asn Glu Glu Tyr Glu
94      225      230      235      240
96 cgt gtc gaa gcc ctg gac aac ggg gtg atc ctg cac ctg aag tcc ggc 768
97 Arg Val Glu Gly Leu Asp Asn Gly Val Ile Leu His Leu Lys Ser Gly
98      245      250      255
100 aag aag atc aag gac gac gcc ttc ctg tgg agc aac ggc cgt acc agc 816
101 Lys Lys Ile Lys Ala Asp Ala Phe Leu Trp Ser Asn Gly Arg Thr Gly
102      260      265      270
104 aat acc gac aag ctg gcc ctg gag aac atc ggt ctg aag gcc aat agt 864
105 Asn Thr Asp Lys Leu Gly Leu Glu Asn Ile Gly Leu Lys Ala Asn Gly
106      275      280      285
108 cgc gaa cag atc cag gtc gac gag cac tac cgt acc gaa gtc agc aac 912
W--> 109 Arg Gly Gln Ile Gln Val Asp Glu His Tyr Arg Thr Glu Val Ser Xaa
110      290      295      300
112 att tat gcc gct ggt gac gtg atc ggc tgg ccc agc ctg gcc agc gcc 960
113 Ile Tyr Ala Ala Gly Asp Val Ile Gly Trp Pro Ser Leu Ala Ser Ala
114      305      310      315      320
116 gcc tat gac cag ggt cgt tgg gcc gcc gcc agt atc acc gag aac gat 1008
117 Ala Tyr Asp Gln Gly Arg Ser Ala Ala Gly Ser Ile Thr Glu Asn Asp
118      325      330      335
120 agc tgg cgt ttc gtc gac gac ctg ccc acc gcc atc tac acc att ccc 1056
121 Ser Trp Arg Phe Val Asp Asp Val Pro Thr Gly Ile Tyr Thr Ile Pro
122      340      345      350
124 gag atc agt tgg gtc gcc aag acc gag cgc gaa ctg acc cag gcc aag 1104
125 Glu Ile Ser Ser Val Gly Lys Thr Glu Arg Glu Leu Thr Gln Ala Lys
126      355      360      365

```

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:12:10

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

```

128 gtt ccc tac gag gtc ggc aag ggt ttc ttc aag ggc atg ggt cgg gca 1152
129 Val Pro Tyr Glu Val Gly Lys Ala Phe Phe Lys Gly Met Ala Arg Ala
130      375      375      390
131 aag atc gcc gtc gag aag ggc ggc atg atg aag atc ctc ttt cac tgc 1290
132 Gln Ile Ala Val Glu Lys Ala Gly Met Leu Lys Ile Leu Phe His Arg
133 385      390      395      400
134 paa acc cta gaa atc ctc ggc gtc cac tgc ttc ggt tat cag gct tgg 1248
135 Glu Thr Leu Glu Ile Leu Gly Val His Cys Phe Gly Tyr Gln Ala Ser
136      405      410      415
137 paa atc gtc cat atc ggc gag ggc atc atg aac cag aag ggc gag ggc 1296
138 Glu Ile Val His Ile Gly Gln Ala Ile Met Asn Gln Lys Gly Glu Ala
139      420      425      430
140 aat acc ctc aag tat ttc atc aac acc acc ttc aac tac cgg acc atg 1344
141 Asn Thr Leu Lys Tyr Phe Ile Asn Thr Thr Phe Asn Tyr Pro Thr Met
142      435      440      445
143 gcc gag gcc tac gag gtc ggc gcc tac gag ggt ctc aat cgg ctt ttt 1392
144 Ala Glu Ala Tyr Ala Val Ala Ala Tyr Asp Gly Leu Asn Arg Leu Phe
145      450      455      460
146 paa 1395

```

```

147 <10> SEQ ID NO: 1
148 <11> LENGTH: 464
149 <12> TYPE: PRT
150 <13> ORGANISM: Anotobacter vinelandii
151 <14> FEATURE:
152 <15> NAME/KEY: misc_feature
153 <16> LOCATION: (5)..(5)
154 <17> OTHER INFORMATION: The 'Xaa' at location 5 stands for Lys, or Asn.
155 <18> FEATURE:
156 <19> NAME/KEY: misc_feature
157 <20> LOCATION: (304)..(304)
158 <21> OTHER INFORMATION: The 'Xaa' at location 304 stands for Lys, or Asn.
159 <22> SEQUENCE: 2

```

```

W--> 172 Met Ala Val Tyr Xaa Tyr Asp Val Val Val Ile Gly Thr Gly Pro Ala
173      1      5      11      15
174 Gly Gln Gly Ala Ala Met Asn Ala Val Lys Ala Gly Arg Lys Val Ala
175      20      25      30
176 Val Val Asp Asp Arg Pro Gln Val Gly Gly Asn Cys Thr His Leu Gly
177      35      40      45
178 Thr Ile Pro Ser Lys Ala Leu Arg His Ser Val Arg Gln Ile Met Gln
179      50      55      60
180 Tyr Asn Asn Asn Pro Leu Phe Arg Gln Ile Gly Glu Pro Arg Trp Phe
181      65      70      75      80
182 Ser Phe Ala Asp Val Leu Lys Ser Ala Glu Gln Val Ile Ala Lys Gln
183      85      90      95
184 Val Ser Ser Arg Thr Gly Tyr Tyr Ala Arg Asn Arg Ile Asp Thr Phe
185      100      105      110
186 Phe Gly Thr Ala Ser Phe Cys Asp Glu His Thr Ile Glu Val Val His
187      115      120      125
188 Leu Asn Gly Met Val Glu Thr Leu Val Ala Lys Gln Phe Val Ile Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,414A

DATE: 01/31/2003

TIME: 11:12:11

Input Set : A:\NIELSEN4.txt

Output Set : N:\CRF4\01312003\I743414A.raw

```

197      197      198      140
198 Thr Gly Ser Arg Pro Tyr Arg Pro Ala Asp Val Asp Phe Thr His Pro
200 145      150      155      160
201 Arg Ile Tyr Asp Ser Asp Thr Ile Leu Ser Leu Gly His Thr Pro Arg
203      165      170      175
204 Arg Leu Ile Ile Tyr Gly Ala Gly Val Ile Gly Cys Glu Tyr Ala Ser
206      180      185      190
208 Ile Phe Ser Gly Leu Gly Val Leu Val Asp Leu Ile Asp Asn Arg Asp
209      195      200      205
211 Gln Leu Leu Ser Phe Leu Asp Asp Glu Ile Ser Asp Ser Leu Ser Tyr
213      210      215      220
214 His Leu Arg Asn Asn Asn Val Leu Ile Arg His Asn Glu Glu Tyr Glu
215 225      230      235      240
217 Arg Val Glu Gly Leu Asp Asn Gly Val Ile Leu His Leu Lys Ser Gly
219      245      250      255
221 Lys Lys Ile Lys Ala Asp Ala Phe Leu Trp Ser Asn Gly Arg Thr Gly
223      260      265      270
225 Asn Thr Asp Lys Leu Gly Leu Glu Asn Ile Gly Leu Lys Ala Asn Gly
227      275      280      285
W--> 226 Arg Gly Gln Ile Gln Val Asp Glu His Tyr Arg Thr Glu Val Ser Xaa
228      290      295      300
230 Ile Tyr Ala Ala Gly Asp Val Ile Gly Trp Pro Ser Leu Ala Ser Ala
232 305      310      315      320
234 Ala Tyr Asp Gln Gly Arg Ser Ala Ala Gly Ser Ile Thr Glu Asn Asp
236      325      330      335
238 Ser Trp Arg Phe Val Asp Asp Val Pro Thr Gly Ile Tyr Thr Ile Pro
240      340      345      350
242 Glu Ile Ser Ser Val Gly Lys Thr Glu Arg Glu Leu Thr Gln Ala Lys
244      355      360      365
246 Val Pro Tyr Gln Val Gly Lys Ala Phe Phe Lys Gly Met Ala Arg Ala
248      370      375      380
250 Gln Ile Ala Val Glu Lys Ala Gly Met Leu Lys Ile Leu Phe His Arg
252 385      390      395      400
254 Glu Thr Leu Gln Ile Leu Gly Val His Cys Phe Gly Tyr Gln Ala Ser
256      405      410      415
258 Glu Ile Val His Ile Gly Gln Ala Ile Met Asn Gln Lys Gly Glu Ala
260      420      425      430
262 Asn Thr Leu Lys Tyr Phe Ile Asn Thr Thr Phe Asn Tyr Pro Thr Met
264      435      440      445
266 Ala Glu Ala Tyr Arg Val Ala Ala Tyr Asp Gly Leu Asn Arg Leu Phe
268      450      455      460
261 <10> SEQ ID NO: 3
262 <11> LENGTH: 52
263 <12> TYPE: DNA
264 <13> ORGANISM: Artificial Sequence
265 <20> FEATURE:
267 <23> OTHER INFORMATION: Description of Artificial Sequence: Primer
268 <400> SEQUENCE: 3
270 attcatgat gaattctat ttatggtcca attctttact gaactgatta ca      52

```

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:19

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

```

273 <210> SEQ ID NO: 4
274 <211> LENGTH: 11
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence
278 <220> FEATURE:
279 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
281 <400> SEQUENCE: 4
282 ggcactact catatcaaag caticctctcg ctggtaatt ttctgtctc ttgtctatca 60
283 gacttagaa a                                     71
286 <210> SEQ ID NO: 5
287 <211> LENGTH: 14
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
294 <400> SEQUENCE: 5
295 ggaattcgc actctcaaaa agag                                     24
298 <210> SEQ ID NO: 6
299 <211> LENGTH: 14
300 <212> TYPE: DNA
301 <213> ORGANISM: Artificial Sequence
303 <220> FEATURE:
304 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
306 <400> SEQUENCE: 6
307 gggagagac ttctagaatg ctttttgata acaaaaaat                                     38
310 <210> SEQ ID NO: 7
311 <211> LENGTH: 18
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
318 <400> SEQUENCE: 7
319 gggagagac tccggcgaga gccataaacg ttaacaaa                                     48
322 <210> SEQ ID NO: 8
323 <211> LENGTH: 19
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
330 <400> SEQUENCE: 8
331 gtcacacac aaggtccta                                     19
334 <210> SEQ ID NO: 9
335 <211> LENGTH: 38
336 <212> TYPE: DNA
337 <213> ORGANISM: Artificial Sequence
339 <220> FEATURE:
340 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
342 <400> SEQUENCE: 9
343 gggggggatc ctctagaatg ccagtgtga aatcagac                                     38

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/743,414A

DATE: 1/31/2003
TIME: 11:52:11

Input Set : A:\NIELSEN4.txt
Output Set : N:\CRF4\01312003\I743414A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 5,304

Seq#:2; Xaa Pos. 5,304

VERIFICATION SUMMARY

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:11

Input Set : A:\NIELSEN4.txt

Output Set : N:\CRF4\01312003\I743414A.raw

L:35 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1, line#:33
L:37 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:48
L:109 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:912
L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:226 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:288